How Certain are Dutch Households about Future Income? An Empirical Analysis

Das, J.W.M.; Donkers, A.C.D.

Publication date:
1997

How certain are Dutch households about future income? An empirical analysis

Marcel Das and Bas Donkers [1]

Dept. of Econometrics and CentER

Tilburg University

April 1997

JEL-classification: C81, D12, D84

Key words: subjective information, income expectations, income uncertainty

Abstract

The growing literature on precautionary saving clearly indicates the need for measurement of income uncertainty. In this paper we empirically analyze subjective income uncertainty in the Netherlands. Data come from the Dutch VSB panel. We measure income uncertainty directly by asking questions on expected household income in the next twelve months. First, we describe our data and compare a measure of income uncertainty with corresponding studies conducted in the US and Italy. Second, we investigate the relationship between the measure of income uncertainty and some household characteristics. Controlling for information on expected changes, we find strong relationships between labor-market characteristics and the subjective income uncertainty as reported by the heads of households.

[1] We are grateful to Rob Alessie, Bertrand Melenberg, Arie Kapteyn, Arthur van Soest, and Peter Wakker for valuable comments. Any remaining errors are ours.
Financial support by the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged by the second author. In this paper we use data collected under the VSB-CentER Savings Project. These data are commonly referred to as the VSB panel.

Please address correspondence to Tilburg University, Department of Econometrics and CentER, P.O. Box 90153, 5000 LE Tilburg, the Netherlands (e-mail: DAS@kub.nl)

1 Introduction

In the dynamic process of household decision making, expectations about the future play a central role. Common versions of the Life Cycle and Permanent Income Hypothesis models assert that current consumption depends not only on current wealth, income and preferences, but also on the individual's or household's subjective distribution of future income. On the basis of an empirical study, Carroll (1994) finds that, for fixed permanent income, current consumption is not influenced by predictable changes in future income. However, future income *uncertainty* has an important effect: consumers facing greater income uncertainty consume less.

In the literature on precautionary saving (cf. Kimball, 1990), several papers have addressed the theoretical result that consumers postpone their consumption when income becomes more risky. See e.g. Guiso et al. (1992), Lusardi (1993), and Banks et al. (1995). Portfolio decisions may also be affected by income uncertainty (Kimball, 1993). At an empirical level, this is illustrated by Guiso et al. (1996): the portfolio share of risky assets is inversely related to income risk.

Most of the empirical studies in which income uncertainty is involved face the problem of measuring the (subjective) uncertainty of future income. Simulation studies and all kinds of proxies for uncertainty of future income are used. But, as noted by Guiso et al. (1992), both approaches have serious drawbacks. For example, simulations do not test whether people actually respond to risk as predicted by the theoretical models. At the cross-sectional level, moreover, indicators for risk are subject to a problem of self-selection.[1] In addition, studies that rely on realized incomes require certain assumptions about the process that generates income in order to construct a proxy for uncertainty of future income. Given the unobservable nature of households' subjective assessments of

[1]Households in risky categories may have chosen to belong to that category simply because they are less risk-averse. Occupational dummies to classify households in different risk categories then give a wrong indication for perceived income uncertainty.

specific risks, Guiso et al. argue that there is no alternative as to rely upon direct measurement of households' perceived uncertainty. Recent work on the subjective measurement of income expectations has indicated that survey data can provide useful information (see e.g. Dominitz and Manski, 1994, and Das and Van Soest, 1996, 1997). The latter show that the relation between answers to subjective survey questions on income expectations and various background variables are rather robust over time and of the expected sign.

This paper focusses on some measures of uncertainty of future income based on subjective data. In attempting to explain relationships between the subjective uncertainty and some household characteristics, our approach follows the study by Dominitz and Manski (1994, DM94 in the sequel), who collected data on the one-year-ahead income expectations on the household level of members of American households [Survey of Economic Expectations (SEE)]. Based on the answers of 437 respondents, they find a substantial variation in income uncertainty. We will use the third wave of a Dutch panel data set: the VSB panel[2] (in this wave the questions we will use are asked for the first time). The panel contains information on more than 2500 households and consists of two subpanels. One is designed to be representative of the whole Dutch population and the other is a random sample from households in the upper 10% of the income distribution in the Netherlands. All participating households have been provided with a personal computer and answer the survey questions directly on their PC. No personal interviews are held.

DM94 compare their study with Guiso et al. (1992) who investigate income uncertainty in Italy. Although aware of the fact that the two survey methods were not the same, they argue that it is tempting to conclude that US households perceive far more income uncertainty than do those in Italy. Results based upon our survey data suggest that also in the case of Dutch households the perceived income uncertainty is lower than in case of US households.

[2]The VSB panel has been supported by the VSB Foundation, which explains its name.

The outline of the paper is as follows. Section 2 discusses the questions posed in the VSB panel to elicit information about subjective income uncertainty. In particular, we will examine two different types of questions: one that is qualitative in nature and a second question that elicits information on income uncertainty in a quantitative way. Section 3 will present the data. Here, the answers to the quantitative questions will be used to derive some measures of income uncertainty that will be compared with those obtained in previous studies. The quantitative measure is also briefly compared with a qualitative measure of income uncertainty. Section 4 estimates a regression model for the location and scale of the subjective income distribution. Section 5 concludes.

2 Data from the VSB panel

The VSB panel started in 1993. The survey method is completely computerized. Each household is provided with a personal computer with a modem. Questions and answers are transferred via the computer. If the respondent has questions or problems he may call a helpdesk.

The data that we will use are taken from the third wave of the panel. These data were collected in 1995 and contain information about 2574 heads of households.[3] The VSB panel consists of two parts. One is designed to be representative for the whole Dutch population, the other one is a random sample of households in the upper 10% of the income distribution in the Netherlands. The information in the data set can be divided into seven parts: household characteristics, housing, labor-market status and pension entitlements, health, income, assets and liabilities, and economic and psychological concepts. Our analysis draws heavily upon the parts concerned with household characteristics, income, and economic and psychological concepts.

[3]The data set also contains information on other household members, but here we focus on heads of households.

The 1995 wave contains two blocks of questions related to the measurement of subjective income uncertainty. The first one consists of qualitative questions and the second one consists of quantitative questions similar to those in DM94. We will discuss both types of questions in the next two subsections.

2.1 Qualitative measurement of uncertainty

All questions in the survey concerning future income are on the household level. Respondents are asked what will happen to their net household income in the next twelve months.[4] First they are asked to indicate whether it will decrease, stay the same or increase. After that, when they indicate they expect a change in income, they are asked by which percentage they think their net household income will change. These questions refer to the location of their distribution of future income and are unrelated with uncertainty. Seven questions related to uncertainty about future income follow directly after the previously mentioned questions. First, respondents are asked how probable an income increase of more than 15% is. They can answer on a seven-point scale ranging from very unlikely to very likely.[5] The same type of question is asked for an increase of between 10% and 15%, between 5% and 10%, no change, a decrease of between 5% and 10%, a decrease of between 10% and 15% and a decrease of more than 15%.

An extensive literature exists on quantifying verbal probability questions. See, among others, Reagan et al. (1989), and Mosteller and Youtz (1990). The former examine the meanings of 18 verbal probability expressions and conclude that some areas of the probability range are not so well captured. The latter try to quantify the meanings of 52 qualitative probabilistic expressions. In a comment on this paper, Kadane (1990) argues that significantly fewer than 52 words are needed. He summarizes the findings of

[4] For the precise wording of the questions, see Appendix A.

[5] Respondents get some information on how to interpret the scale. However, only the end-points of the seven-point scale have a verbal label.

Mosteller and Youtz into eleven verbal descriptions that cover the whole range of possible probabilities.

In this literature, some authors prefer verbal, nonnumerical terms for communicating uncertain opinions. Wallsten et al. (1986) argue that most people feel that they better understand words than numbers. On the contrary, Beyth-Marom (1982) highlights the communication problems caused by verbal probability expressions. In addition to the better communication achieved by numerical expressions, another advantage is the possible application of various quantitative methods.

This paper focusses on the quantitative expressions, since to the best of our knowledge, no work has been done on deriving a (characteristic of the) subjective probability distribution from verbal questions. The end of Section 3 briefly compares a measure of uncertainty derived from the qualitative questions mentioned above and a measure of uncertainty derived from the quantitative questions.

2.2 Quantitative measurement of uncertainty

The qualitative type of questions mentioned in the previous subsection are asked in each wave of the VSB panel. Since 1995 there are also questions in the panel that try to elicit the subjective distribution of future income in a quantitative way. First, the respondents are asked about the range in which their household income will fall in the next twelve months. The precise wording of the questions is as follows:

> *What do you think is the LOWEST level your net household income could possibly be over the next twelve months?*

and

> *What do you think is the HIGHEST level your net household income could possibly be over the next twelve months?*

After answering these two questions, the respondents are asked to evaluate the probability (in percentage terms) with which their household income will fall below a certain level. Four questions of this type are asked, where the levels referred to in these questions are evenly spread over the interval ranging from the household's reported lowest possible income to the highest possible income. [6] The precise wording of the question is as follows:

> *How large do you think is the probability that the total net income of your household in the next twelve months will be below level$_k$? Please give a number between 0 and 100.*

The answers to these questions will be denoted by $\text{PRO}_1, \ldots, \text{PRO}_4$ and correspond to values of the subjective distribution function of next year's household income.

Similar questions are used by DM94 to investigate income expectations. The first difference between our data and the data from the SEE used by DM94 is that the levels to which the questions in our data refer are evenly spread over the range of possible realizations of next year's household income, while the levels in the SEE questions are taken from a given sequence. Given the validity of the lowest and highest possible realizations, there will be no anchoring effect present in our data.[7] Given the midpoint between the lowest and highest possible income, DM94 select four values from a predetermined sequence of income thresholds in such a way that two thresholds are below and two thresholds are above the midpoint. This way of selecting thresholds avoids some anchoring problems, although it does not remove them completely. Respondents who are quite uncertain about their household income will see reasonable values for the thresholds, but if the head

[6] Evenly spread means that the level in question k ($k = 1, \ldots, 4$) is equal to: *lowest possible income* + $0.2k$ (*highest possible income - lowest possible income*).

[7] Anchoring means that a respondent adapts his beliefs to the questions that are asked. If a respondent believes that the household income will never be below say Dfl 40,000 he might be induced to give positive probabilities to outcomes below this value. This can be the case if, for example, the levels that are referred to are all below this level of Dfl 40,000. The reasoning of the respondent in this case is that his beliefs might be wrong since the researcher seems to be interested in these low outcomes. The respondent might think that these values are objectively reasonable.

of household is certain about the household income in the next twelve months (say the difference between highest and lowest possible income is Dfl 2,000), he will face rather low and high values for the thresholds, which might in turn induce him to spread his subjective density more widely.

The second difference between our data and the data from the SEE is that in the SEE, if a respondent gave an answer that was incompatible with the previous ones, this inconsistency was mentioned to the respondent. A new answer was then given. This way of questioning results in a higher fraction of valid answers and will be pursued in the next wave of the VSB panel. For the current wave we will have to ignore the respondents who provided an inconsistent sequence of probabilities.

3 Measurement of subjective income uncertainty

For the measurement of the subjective income uncertainty we will use the quantitative questions described in Section 2.2. These questions can be found in the income part of the panel. The 1995 wave of the panel consists of 2574 heads of households.[8] Only 1614 of them answer affirmatively a question on whether or not they have an idea about their household's income in the past year. These heads of households all answer the question what the household's lowest and highest possible income for the next year will be. After deleting households with extremely low values for their income and a few households giving a higher value for the lowest possible income than for the highest possible income, 1504 households remain with observed lowest and highest possible income levels for the next twelve months.

Following the questions on lowest and highest possible incomes, the heads of households are asked to evaluate the probability with which their household income will fall

[8] The representative and the high income part of panel are combined.

below a certain level (see Section 2.2). Four questions of this type are asked, and in theory, the given probabilities should result in a non-decreasing sequence of answers. This is not true for 220 of the heads of households, while two heads of households do not answer the questions. In addition to the questions from the income part of the questionnaire, also some questions from the economic and psychological part will be used. These questions are related to realized and expected income changes of the household's income (see Section 2.1). Due to some missing observations, our final data set consists of 1127 heads of households with completely observed information from both parts of the questionnaire.

Some descriptive statistics concerning both the lowest and highest possible income and the probabilities (in percentages) are given in Table 1. We distinguish between the representative and high-income part of the panel to see whether there are systematic differences.

The numbers in Table 1 indicate that there is substantial variation in the respondents' answers to $\text{PRO}_1, \ldots, \text{PRO}_4$. Further, we see that the answers to the probability questions are similar for the representative and high-income panel, whereas the stated possible incomes are higher for the high-income panel, as could be expected. This suggests that if we condition on income, we need not distinguish between the two parts of the panel.

In choosing a measure of income uncertainty, we will follow DM94. They use the interquartile range of the subjective distribution of next year's income as a measure of income uncertainty. To calculate this interquartile range, we specify a distribution function known up to a parameter (vector) θ and then estimate θ using our data (see also DM94). That is,

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \sum_{k=1}^{4} \left(\frac{\text{PRO}_k}{100} - F(\text{level}_k; \theta)\right)^2, \tag{1}$$

where $F(.;\theta)$ is a distribution function with unknown parameter θ. The parameter θ can then be estimated with Non-Linear Least Squares.

DM94 chose a lognormal distribution with a two-dimensional parameter vector θ: the

median (to characterize the central tendency) and the interquartile range (to characterize its dispersion). Estimation is not possible for households with at least three times a value of zero or one. The best fitting distribution in that case is a degenerate distribution with all mass at level k for which the corresponding PRO_k is unequal to zero or one.

Table 1 : Descriptive statistics for the answers to the quantitative questions for the representative and high-income part of the panel

	Lowest Income	Highest Income	PRO_1	PRO_2	PRO_3	PRO_4
Representative part of panel; 805 observations						
Minimum	3,000	5,000	0	0	0	0
1st Quartile	26,244	31,200	1	10	20	40
Median	40,000	45,000	10	25	50	70
3rd Quartile	54,000	60,000	25	50	75	90
Maximum	330,000	360,000	100	100	100	100
Mean	41,488	48,214	19.4	32.3	49.3	61.8
Std. Dev.	25,367	31,619	24.2	28.2	31.2	31.4
High-income part of panel; 322 observations						
Minimum	3,000	5,000	0	0	0	0
1st Quartile	40,000	55,000	0	10	20	40
Median	70,000	80,000	10	25	50	70
3rd Quartile	86,000	100,000	25	50	70	90
Maximum	300,000	800,000	100	100	100	100
Mean	64,363	77,547	17.3	29.6	46.0	61.4
Std. Dev.	39,910	61,788	23.3	26.8	31.5	32.3

Note: 205 respondents gave the same answer on the questions for the lowest and highest possible income. For these observations, the values for $\text{PRO}_1, \ldots, \text{PRO}_4$ are not determined, so they are not used in the last four columns.

DM94 compare their results with another study using survey data on future income expectations: a biennial survey of the Bank of Italy [the Survey of Household Income and Wealth (SHIW)]. The SHIW elicited points of the subjective probability distributions for

the growth rate of nominal labor earnings and pensions and for the rate of inflation over the next twelve months.[9] Guiso et al. (1992) use the ratio of the standard deviation (σ) to the mean (μ) of the subjective real income distribution to measure subjective earnings uncertainty. Their results, the results of DM94, and our results based on the estimator in (1) are summarized in Table 2.

Table 2 (columns two, three and four) shows that the income uncertainty in the Netherlands, as measured by the coefficient of variation, is between the income uncertainty in Italy and the US. This result suggests that Dutch households perceive more income uncertainty than Italian households do, but households in the US face more income uncertainty than households in the Netherlands. A χ^2-test has been used to test whether the difference in uncertainty between the US and the Netherlands as tabulated in Table 2 is significant. The resulting test statistic is equal to 408, exceeding the critical value of 26.3. It should be mentioned that part of this result might be caused by different survey methods. However, the type of questioning and the estimation procedure in the SEE and in the VSB panel are similar. In that respect, the US and the Dutch results are comparable; it therefore seems safe to conclude that perceived income uncertainty is smaller in the Netherlands than it is in the US.

A disadvantage of using the lognormal distribution is the fact that we do not use explicitly the information on the reported lowest and highest possible income. The lognormal distribution also takes values outside the interval [lowest possible income, highest possible income]. In our case, a substantial part of the total probability mass is out-

[9]The exact wording of the SHIW question on the subjective probability distribution is: *We are interested in knowing your opinion about labor earnings or pensions twelve months from now. Suppose that you have 100 points to be distributed between these intervals* (a table is shown to the person interviewed). *Are there intervals which you definitely exclude? Assign zero points to these intervals. How many points do you assign to each of the remaining intervals?* For this and a similar question on inlation uncertainty the intervals of the table shown to the person interviewed are: $> 25, 20-25, 15-20, 13-15, 10-13, 8-10, 7-8, 6-7, 5-6, 3-5, 0-3, < 0$ percent. In case it is less than zero, the person is asked: *How much less than zero? How many points would you like to assign to this class?* For further details on the Italian SHIW, see Guiso et al. (1992).

side the interval. To give an indication, for almost 30% of all the respondents with a non-degenerate subjective distribution, more than half of the total probability mass lies outside the interval. Moreover, for approximately 20% of all the respondents with a non-degenerate subjective distribution, the median lies outside the interval. This seems unrealistic. The fact that the lognormal distribution gives a good approximation to the distribution of household incomes over the population does not imply that this is also the case for (subjective) income distributions on the household level.

Table 2 : Relative frequency distributions of the variation coefficient of future income

	Italian SHIW	US SEE	Dutch VSB panel	Dutch VSB panel	Dutch VSB panel
			Lognormal	Beta	Interpol.
$\sigma/\mu = 0.000$	0.34	0.20	0.28	0.28	0.18
$\sigma/\mu \leq 0.005$	0.44	0.20	0.30	0.32	0.28
$\sigma/\mu \leq 0.015$	0.70	0.20	0.36	0.44	0.44
$\sigma/\mu \leq 0.025$	0.88	0.20	0.47	0.58	0.58
$\sigma/\mu \leq 0.035$	0.94	0.21	0.55	0.67	0.66
$\sigma/\mu \leq 0.045$	0.99	0.22	0.62	0.75	0.73
$\sigma/\mu \leq 0.065$	1.00	0.24	0.71	0.84	0.82
$\sigma/\mu \leq 0.100$	1.00	0.34	0.81	0.93	0.91
$\sigma/\mu \leq 0.150$	1.00	0.44	0.89	0.96	0.95
$\sigma/\mu \leq 0.200$	1.00	0.53	0.92	0.99	0.97
$\sigma/\mu \leq 0.300$	1.00	0.70	0.96	1.00	0.99
$\sigma/\mu \leq 0.400$	1.00	0.78	0.98	1.00	1.00
$\sigma/\mu \leq 0.500$	1.00	0.85	0.98	1.00	1.00
$\sigma/\mu \leq 1.000$	1.00	0.94	0.99	1.00	1.00
$\sigma/\mu \leq 2.000$	1.00	0.98	1.00	1.00	1.00
$\sigma/\mu \leq 5.000$	1.00	0.99	1.00	1.00	1.00
# observations	2,909	437	982	982	1127

Note: For the Dutch VSB panel, the estimation procedure for the unknown parameter vector in case of the lognormal and Beta distribution does not converge when the respondent gave the same answer to all $\text{PRO}_1, \ldots, \text{PRO}_4$. For this reason we could not use all the observations.

We can explicitly use the information on the reported lowest and highest possible incomes by putting all the probability mass on the reported interval. A possible distribution that takes this into account is the Beta distribution. This family of distributions is flexible in that it covers both symmetric and asymmetric distributions.

The effect of estimating a distribution function defined on the reported interval becomes clearer when we look at the fifth column of Table 2. This column displays the variation coefficient of future income in the Netherlands when we use estimates derived from a Beta distribution. We see that the relative frequencies in the Dutch case come closer to the Italian numbers.

When estimating the lognormal or Beta distribution, we cannot use the observations where the respondent gave the same answer to all $\text{PRO}_1, \ldots,\text{PRO}_4$. This means a loss of 145 observations. But all these respondents gave a useful answer to the lowest and highest possible income and therefore provided useful information on their subjective income uncertainty. If we assume that the density of the subjective income distribution is simply (piecewise) uniform over the intervals, we are able to use these observations. In this case, we can obtain the estimated cumulative distribution function by interpolation between the known points 0, $\text{PRO}_1, \ldots,$ PRO_4, and 100. The relative frequency distribution of the variation coefficient in case of the interpolated distribution is presented in the sixth column of Table 2. Only for small values of the variation coefficient do we find differences with column 5. The characteristics such as median or interquartile range are similar in case of interpolation compared to the estimated Beta distribution. In all further analyses we will use the characteristics of the piecewise uniform distribution function.

The rank correlation between IQR and MED is 0.43 and highly significant. It would be interesting to know what the relationship is between the expected level of income and subjective income uncertainty. In the case where IQR is proportional to MED, the relative income uncertainty (IQR/MED) is constant. Using our data, we (nonparametrically)

regress the quotient IQR/MED on MED. The result is presented in Figure 1. Together with the estimated functional relationship between IQR/MED and MED, we present 95% uniform confidence bounds. [10]

Figure 1 : Nonparametric regression of relative subjective income uncertainty (IQR/MED) on the subjective median of future income (MED). The dashed lines are 95% uniform confidence bands.

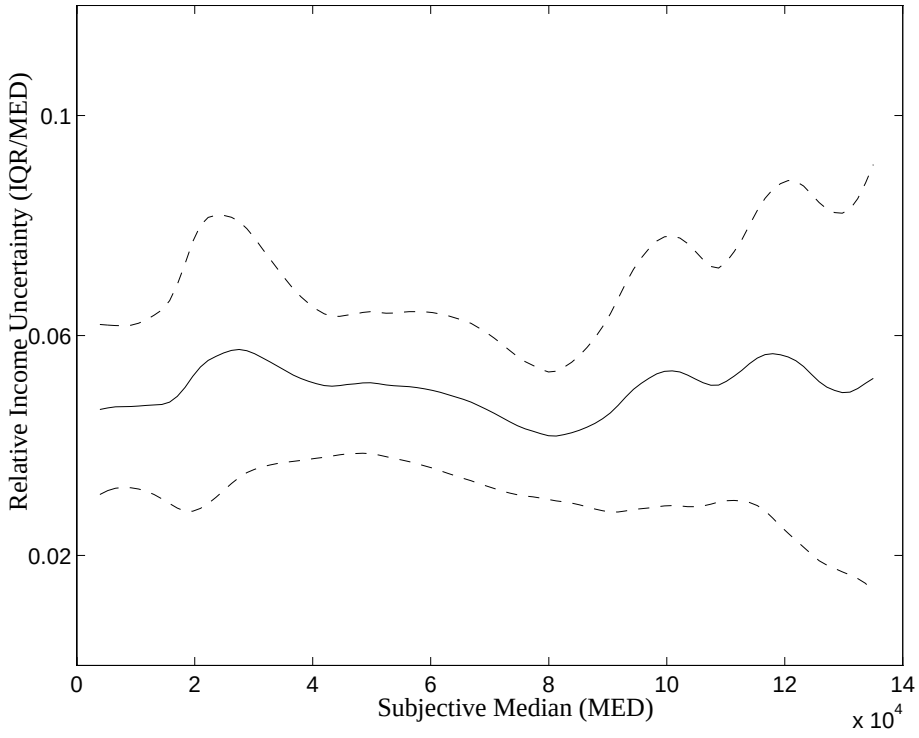

Figure 1 shows that the median of the subjective income distribution has no significant effect on relative income uncertainty as perceived by the head of household. This implies that households that expect a higher income next year do not perceive a greater or

[10] We use the quartic kernel and a bandwidth equal to $1,5\times10^4$. For details on nonparametric regression, see e.g. Härdle and Linton (1994).

smaller *relative* uncertainty than others do. In studies by Skinner (1988), Zeldes (1989), and Carroll (1992), the household's subjective IQR is also proportional to the median. However, these studies rely on realizations and on a log-normality assumption,[11] while our conclusion is based on survey data on subjective income expectations.

Qualitative versus Quantitative measurement

As we already mentioned in Section 2.1, we also have some qualitative questions related to the expectations on household income in the next twelve months. These questions are related to changes in household income in relation to income in the past twelve months (for the precise formulation of the questions, see Appendix A).

We constructed probabilities for the categories of income changes by assigning weights to the different answering categories. Since for a given question a higher category corresponds to a higher likelihood, we assign increasing weights to the categories for each question. A general way to do this would be as follows:

$$W_j = \alpha_j + \beta_j * \text{number}_j, \text{ with}$$
$$j = 1 \text{ (Increase of more than 15\%)} , \ldots , 7 \text{ (Decrease of more than 15\%)} .$$

Here W_j is the weight assigned to the income change for category j if number$_j$ was the number for the answer category for income change j, with number$_j$ in {1(*highly unlikely*),...,7(*highly likely*)}. The reason the weights are modelled this way is mainly because the probability related to *highly likely* does not necessarily have to be equal to 7 times the probability of *highly unlikely*. Since there are no verbal clarifications for the answers between the extremes 1 (*highly unlikely*) and 7 (*highly likely*), we see no reason why we should not assume equal increases for the probabilities corresponding to

[11] Carroll (1992) superimposes a 0.005 chance of receiving no income at all. As already mentioned by DM94, this slight modification of the log-normality assumption has a negligible effect on the median and IQR of the subjective income distribution.

the answer within the range *highly unlikely*, ..., *highly likely*. All we know is that $\beta_j > 0$ and $\alpha_j + \beta_j \geq 0$.

To obtain probabilities from these weights, we simply normalized them to sum to one, that is

$$P_j = \frac{W_j}{\sum_{k=1}^{7} W_k}.$$

Adding these probabilities yields six points on the cumulative distribution function for the expected income changes. We can derive a median change and the interquartile range of income changes by interpolation between the known points of the cumulative distribution function, similar to the procedure used for the quantitative data.

We want to compare the measure of uncertainty obtained from the qualitative questions with that obtained from the quantitative questions. The problem, however, is that the quantitative questions refer to income levels, while the qualitative questions refer to percentage changes from past income. Since we have information on only income classes for the past twelve months' income, we will obtain imprecise results if we use this variable to scale the distribution for income changes to a distribution for expected income levels. When we calculate the ratio of the interquartile range to the median, however, the scale drops out and we obtain the same expression as for the quantitative information.[12] To see whether the qualitative data yields similar outcomes as compared with the quantitative data, we examine the correlation between the ratio of the interquartile range to the median for the two types of data. The rank correlation coefficient is equal to 0.25 and is highly significant.

[12]To be exactly, $\frac{IQR_t(y_{t+1})}{MED_t(y_{t+1})} = \frac{IQR_t(\Delta y_{t+1}/y_t)}{MED_t(\Delta y_{t+1}/y_t)+1}$.

4 Prediction of the subjective measure of income uncertainty.

This section examines how our measure of income uncertainty varies with some household characteristics. A (possible) correlation can yield useful information. First, if we find no correlation at all, this may cast doubt on our measure of income uncertainty based on the subjective data – especially in cases where a relationship between income uncertainty and household characteristics is plausible. Second, if a relationship exists, this information might be useful for studies in which no subjective data are available. In that respect, we try to gain some insight into the way the employment status of the partner affects the income uncertainty of the household.

Before we discuss the results for income uncertainty we will examine the location of the subjective income distribution.

Location

We estimate a simple model for the median of the subjective income distribution (as a measure of location): the same linear specification as used by DM94. We allow for a more flexible age pattern than DM94 and we also distinguish between respondent and spouse with respect to labor-force participation. The exact definitions of the explanatory variables can be found in Appendix B.1. We use LAD estimation to make our estimates robust to outliers and bootstrapping to calculate the asymptotic covariance matrix. The reported standard errors are corrected for potential heteroskedasticity. Table 3 presents the estimation results.

The first column in Table 3 shows that the household income in the past twelve months is a dominant predictor for the expected household income in the next twelve months.

A striking result is that the estimated coefficient is almost the same as found by DM94. The best linear prediction of the location measure of the subjective income distribution increases 834 Dutch guilders with every one thousand Dfl. increase of past household income.

Table 3 : Estimation results for the median

DEPENDENT VARIABLE: MEDIAN (in thousands of Dfl.)				
	without interactions		with interactions	
Constant	7.58	(4.3)	10.7	(4.3)
PastInc	0.834	(0.021)	0.813	(0.036)
PastInc×DumWork			0.101	(0.045)
PastInc×DumWorkP			-0.115	(0.042)
DumWork	2.34	(0.74)	-2.11	(1.7)
DumWorkP	-1.84	(0.82)	3.44	(2.0)
DumUnem	-2.08	(0.79)	-1.79	(1.0)
DumUnemP	-0.277	(1.9)	-0.791	(1.5)
DumFemale	-0.969	(0.59)	-1.31	(0.73)
DumPartner	1.53	(0.76)	1.00	(0.87)
Age/10	-1.36	(1.4)	-1.79	(1.6)
$Age^2/100$	0.135	(0.13)	0.162	(0.15)
DumEdu2	0.772	(0.80)	0.210	(1.1)
DumEdu3	0.431	(0.89)	0.122	(1.2)
DumEdu4	1.58	(1.1)	1.54	(1.2)
DumEdu5	2.34	(1.2)	1.89	(1.7)
DumStartW	0.994	(1.9)	0.232	(1.7)
DumStopW	-4.57	(2.0)	-5.10	(2.1)
Average Abs. Dev	15.8		15.7	

Note: Standard errors are in parentheses.

Heads of households with a higher level of education expect a higher level of income in the next twelve months. However, the joint hypothesis whether all dummy variables corresponding to the level of education are equal to zero cannot be rejected (significance probability of 0.40).

The first column of Table 3 shows also that differences exist between head of household

and partner in the effect of labor-market status on expected income. DM94 consider only the aggregate effect of labor force participation by respondent and spouse. They find no significant influence. Here we see, for example, that if the head of household has a job and a partner is present in the household, the difference in the median between a working and non-working partner is significant and almost Dfl 2,000 (*ceteris paribus*).

The negative sign of the variable DumWorkP might be explained by the type of jobs (and the corresponding salary) partners have. This is best illustrated when we allow household income to interact with the employment dummies for head and partner. The resulting estimates are presented in the second column of Table 3. When we consider a household with a working head and a non-working partner, the coefficient on household income is equal to 0.914. For a household with a working head and working partner, this coefficient is equal to 0.799. This suggests that last year's household income is less dominant in predicting next year's household income when the partner is working. Note that these results are conditional on whether or not the head expects some household member to stop working. This expectation exerts a strong negative effect. The effect of a member in the household who is expected to start working is smaller and insignificant.

The above explanation for the smaller part of last year's household income which is carried over into expectations for the next year also suggests that a household with working head and partner faces more income uncertainty than does a household with working head and non-working partner. This issue will be addressed when we move on to income uncertainty.

Absolute income uncertainty

As mentioned before, we use the InterQuartile Range (IQR) as a measure for income uncertainty. The IQR is a measure for absolute income uncertainty (that is, a guilder more is the same for all households, independent of the level of their income). The end of this section will also address relative income uncertainty.

We use the same model as in the analysis of the median. Instead of using the dummy variables corresponding to start/stop working (which proved to be insignificant), we incorporate some variables referring to expectations about income changes in the past and future. The variable PrevΔInc denotes the subjective change in household income in the last twelve months, and the variable ExpΔInc refers to the expected income change in the next twelve months (both variables are in percentage terms). The estimation results appear in Table 4.

The first column of Table 4 shows that the IQR depends significantly on income in the last twelve months, but the effect is small if we compare it with the results obtained by DM94 for the US. The difference in magnitude is more than tenfold (and confidence intervals do not overlap). This is, of course, related to the earlier finding that Americans perceive far more income uncertainty than Dutch heads of households do.

Furthermore, we find, unlike DM94, a positive effect of a working partner on income uncertainty. Income uncertainty is even higher when the partner is unemployed and searching for a job. A female head of household perceives less income uncertainty than a male, as is shown by the coefficient corresponding to DumFemale being significantly negative.

We included a quadratic age pattern. The estimated coefficients are highly significant. Absolute income uncertainty decreases with age until the age of retirement. Although DM94 don't include a quadratic term, they also find a negative relationship between

income uncertainty (as measured by IQR) and age. The education level has no effect, as is shown by a joint test on the coefficients for the dummy variables (significance probability of 0.62).

Table 4 : Estimation results for the interquartile range

DEPENDENT VARIABLE: IQR (in hundreds of Dfl.)				
	without income change variables		with income change variables	
Constant	47.7	(5.5)	35.5	(10.0)
PastInc	0.123	(0.016)	0.116	(0.042)
DumWork	-1.65	(0.96)	3.42	(1.6)
DumWorkP	2.80	(1.2)	-0.922	(1.8)
DumUnem	0.992	(1.1)	0.879	(2.9)
DumUnemP	11.9	(1.3)	11.5	(2.3)
DumFemale	-2.56	(1.2)	-2.14	(1.2)
DumPartner	-1.30	(1.2)	-0.870	(1.2)
Age/10	-15.1	(2.1)	-11.9	(3.3)
$Age^2/100$	1.15	(0.21)	0.925	(0.28)
DumEdu2	0.570	(2.0)	1.01	(0.91)
DumEdu3	1.06	(1.9)	1.33	(1.3)
DumEdu4	-0.745	(1.9)	-0.389	(0.95)
DumEdu5	0.905	(2.0)	1.20	(1.7)
PrevΔInc			0.0661	(0.17)
|PrevΔInc|			0.125	(0.16)
ExpΔInc			0.105	(0.14)
|ExpΔInc|			0.373	(0.18)
Average Abs. Dev.	2.42		2.40	

Note: Standard errors are in parentheses.

The second column of Table 4 shows the estimation results after we included expectations and perceived realizations of income changes. It turns out that only the absolute value of the expected income change (ExpΔInc) has a significant influence on income uncertainty: the larger the expected change, the more uncertain a head of household is about future income. We included both the expected income change and its absolute value to see whether an expected increase in household income has a different effect than an

expected decrease in household income. This, however, makes no difference. Past income changes have no significant effect.

Relative income uncertainty

The IQR is a measure of income uncertainty that does not take into account the level of income at which the variation in income takes place. This section will examine a measure of relative uncertainty of next year's income by taking the ratio of IQR to MED as our variable of interest. This measure looks at income changes as relative deviations from the median. Estimation results are presented in Table 5.

Results in the first column of Table 5 reveal that household income in the past twelve months has a significant positive effect on the relative income uncertainty, although we could not reject proportionality between IQR and MED (see Figure 1). Note, however, that when the household income is (*ceteris paribus*) Dfl 10,000 higher, the best linear prediction of the relative income uncertainty increases with less than 0.2%.[13]

When we look at the labor-market status variables for head and partner, we see that if a partner has a job, this does not influence relative income uncertainty, whereas the fact that the head of household has a job increases relative income uncertainty by almost one percentage point. The unemployment dummies for head and partner are of the same order of magnitude. (Note, however, that DumUnemP is significant and DumUnem is insignificant.) A test on the joint significance of the dummy variables corresponding to the level of education indicates that there exist differences between education levels (significance probability is equal to 0.03).

When we include some characteristics of past and expected income changes, we obtain the results presented in the second column of Table 5. Again we see that only the absolute

[13]We also included a quadratic term in past income, but this did not change the results, with the quadratic term being insignificant.

magnitude of expected income changes influences income uncertainty in a positive way. The effects of the other variables are the same as in the first column. Only the variable DumWork is no longer significant.

Table 5 : Estimation results for relative income uncertainty

DEPENDENT VARIABLE: 100*(IQR/MED)				
Constant	10.9	(2.0)	9.07	(2.6)
PastInc	0.0145	(0.0065)	0.0128	(0.0048)
DumWork	0.738	(0.21)	0.716	(0.40)
DumWorkP	-0.0852	(0.32)	0.0804	(0.40)
DumUnem	1.27	(0.65)	1.08	(0.61)
DumUnemP	1.78	(0.37)	1.45	(0.57)
DumFemale	-0.786	(0.35)	-0.731	(0.23)
DumPartner	-0.450	(0.42)	-0.451	(0.32)
Age/10	-3.50	(0.62)	-2.91	(0.82)
$Age^2/100$	0.280	(0.052)	0.235	(0.068)
DumEdu2	0.525	(0.32)	0.456	(0.27)
DumEdu3	0.603	(0.40)	0.559	(0.38)
DumEdu4	0.177	(0.26)	0.162	(0.29)
DumEdu5	0.713	(0.43)	0.651	(0.41)
PrevΔInc			0.0222	(0.040)
|PrevΔInc|			0.0321	(0.035)
ExpΔInc			0.0595	(0.047)
|ExpΔInc|			0.0984	(0.035)
Average Abs. Dev.	4.09		4.04	

Note: Standard errors are in parentheses.

Comparing the estimation results in Table 4 and Table 5, we see that the signs of all the significant variables are the same. The age pattern has not changed much: income uncertainty decreases until the age of retirement. The level of education, however, influences relative income uncertainty significantly, while it does not affect absolute income uncertainty. Finally, it should be noted that comparing the magnitude of the effects makes no sense, since we try to explain a different measure of income uncertainty (*absolute* versus *relative*).

5 Conclusions

We have analyzed subjective data on income uncertainty using data from the 1995 wave of the Dutch VSB panel. In the analysis we use questions that elicit the subjective income distribution in a quantitative way. We compare our measure of income uncertainty with corresponding studies conducted in the US and Italy and find that perceived income uncertainty in the US is larger than in the two European countries.

There was also a significant correlation between two different measures of income uncertainty, one measure being derived from qualitative questions, the other from questions with quantitative answers.

The median of the subjective income distribution is used as a measure of the household's income level. We find that the household income in the past twelve months is a dominant predictor for future income. However, last year's household income is less dominant in predicting next year's household income when the partner is working.

We use a measure of future income uncertainty the interquartile range of the subjective income distribution. We distinguish between absolute and relative income uncertainty. For both measures we find that income uncertainty decreases with age until retirement. Furthermore, there is a positive effect of a working partner on income uncertainty. This effect increases when a partner is unemployed and searching for a job.

Results from our analysis suggest that it is worthwhile to use subjective data. This type of data provides useful information and can be used to measure income uncertainty, which is an important aspect in household decision making. A next step would be to explicitly incorporate subjective data on income uncertainty in models explaining household behavior.

Appendix

A. Exact wording of survey questions

"Income" part of questionnaire

On the next screen you will be asked how much, approximately, the TOTAL NET INCOME OF YOUR HOUSEHOLD AS A WHOLE has been over the period 1 January 1994 through 31 December 1994. The total net income of the household means the sum of net incomes of all household members. By net income we mean the income after deduction of taxes, but before making payments for things like rent, mortgages, and the like.

Please indicate about how much the TOTAL NET INCOME OF YOUR HOUSEHOLD was over the period 1 January 1994 through 31 December 1994.
Possible answers: Less than Dfl. 17,500 (1); Dfl. 17,500 - Dfl. 20,000 (2); Dfl. 20,000 - Dfl. 24,000 (3); Dfl. 24,000 - Dfl. 28,000 (4); Dfl. 28,000 - Dfl. 34,000 (5); Dfl. 34,000 - Dfl. 43,000 (6); Dfl. 43,000 - Dfl. 55,000 (7); Dfl. 55,000 - Dfl. 80,000 (8); Dfl. 80,000 - Dfl. 105,000 (9); Dfl. 105,000 - Dfl. 150,000 (10); Dfl. 150,000 or more (11); (Also a *Don't know* category is given.)

We would like to know a bit more about your expectations of total net household income in the next 12 months. What do you think is the LOWEST amount that your total net household income could possibly be over the next 12 months?

The same question is asked for HIGHEST amount of total net household income.

Next we will show you a number of possible amounts of total net household income. Can you indicate for each of these amounts what the probability in percentages is (or number of cases out of 100) that the total net household income in the next 12 months will be LESS than the given amount?

What do you think is the probability that the total net household income in the next 12 months will be less than *[LOWEST + (HIGHEST - LOWEST)*0.2]* ?[14]
Fill in a number between 0 and 100.

This question is repeated for *[LOWEST + (HIGHEST - LOWEST)*0.4], [LOWEST + (HIGHEST - LOWEST)*0.6]*, and *[LOWEST + (HIGHEST - LOWEST)*0.8]*.

[14] Automatically filled in by the computer.

"Economic and psychological concepts" part of questionnaire

The TOTAL NET INCOME OF YOUR HOUSEHOLD consists of the income of all members of the household, after deduction of taxes, taken as the sum total over the past 12 months.

PREVIOUS INCOME CHANGE:

Compared to about one year ago, did the total net income of your household increase, remain about the same, or decrease?
Possible answers: increase (1), remain about the same (2), and decrease (3).

Only for those who filled in a change: By what PERCENTAGE (approximately) has the total net income of your household increased (decreased)?
(Note: for those who filled in *remain about the same* the income change is set to 0% in the analysis.)

FUTURE INCOME CHANGE:

Do you think, taking into account possible changes within the household, the total net income of your household will increase, remain the same, or decrease, IN THE NEXT 12 MONTHS? *Possible answers: increase (1), remain about the same (2), and decrease (3).*

Only for those who filled in a change: By what PERCENTAGE do you think the total net income of your household will increase (decrease) IN THE NEXT 12 MONTHS?
(Note: for those who filled in *remain about the same* the income change is set to 0% in the analysis.)

We would like to know a bit more about your expectations of the next 12 months. Below we have presented a number of possible changes in income. Please indicate (on the scale given) with any of those changes, how likely you think it is that the total income of your household will change by that percentage IN THE NEXT 12 MONTHS.

A rise in income of more than 15%
Possible answers: Highly unlikely (1), (2), (3), (4), (5), (6), and Highly likely (7). Note that only the endpoints (1) and (7) have a verbal explanation. Also a category *Don't know* is given.

The above question is repeated for a rise in income between 10 and 15%, a rise in income between 5 and 10%, no significant change in income (change not more than 5%), a drop in income between 5 and 10%, a drop in income between 10 and 15%, and a drop in income of more than 15%.

B.1 Reference List of Variables

MED	Median; derived from the interpolated subjective expected income distribution.
IQR	Interquartile range; derived from the interpolated subjective expected income distribution.
PastInc	Midpoint of income bracket that contained the household's income in the past twelve months according to the head of household. Eleven brackets are used (see Appendix A). The variable is measured in thousands of Dutch guilders.
DumWork	Dummy variable: 1 if the head of household has a paid job, 0 otherwise.
DumWorkP	Dummy variable: 1 if the partner has a paid job, 0 otherwise.
DumUnem	Dummy variable: 1 if the head of household is unemployed and searching for a job.
DumUnemP	Dummy variable: 1 if the partner is unemployed and searching for a job.
DumFemale	Dummy variable: 1 if the head of household is female, 0 otherwise.
DumPartner	Dummy variable: 1 if there is a partner present in the household.
Age	Age of the head of household.
DumEdu1..5	Dummy variables for education levels in increasing level of education: DumEdu1: primary education DumEdu2: lower secondary education DumEdu3: higher secondary and intermediate vocational education DumEdu4: higher vocational and pre-university education DumEdu5: university education Reference group is DumEdu1.
DumStartW	Dummy variable: 1 if the head of household expects that household income in the next twelve months will be influenced by the fact that *a* member of the household who is currently not employed will start working, 0 otherwise.
DumStopW	Dummy variable: 1 if the head of household expects that household income in the next twelve months will be influenced by the fact that *a* member of the household who is currently employed will stop working, 0 otherwise.

PrevΔInc Previous change in income in the past twelve months. The variable is measured in percentage terms (see Appendix A).

ExpΔInc Expected change in income in the next twelve months. The variable is measured in percentage terms (see Appendix A).

B.2 Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.
MED	52,514	34,609	3,933	342,000
IQR	2,838	6,222	0	128,000
IQR/MED	0.0515	0.0793	0	0.895
PastInc	62.5	31.1	15	175
DumWork	0.717	0.451	0	1
DumWorkP	0.382	0.486	0	1
DumUnem	0.161	0.367	0	1
DumUnemP	0.0967	0.296	0	1
DumFemale	0.137	0.344	0	1
DumPartner	0.800	0.400	0	1
Age	49.0	13.2	22	88
DumEdu1	0.0481	0.214	0	1
DumEdu2	0.183	0.386	0	1
DumEdu3	0.268	0.443	0	1
DumEdu4	0.305	0.460	0	1
DumEdu5	0.197	0.398	0	1
DumStartW	0.0258	0.158	0	1
DumStopW	0.0506	0.219	0	1
PrevΔInc	0.399	8.06	-80	100
|PrevΔInc|	2.98	7.50	0	100
ExpΔInc	-0.239	8.78	-40	100
|ExpΔInc|	3.03	8.24	0	100

References

Banks, J., R. Blundell, and A. Brugiavini (1995), Income Uncertainty and Consumption Growth, *IFS Working Paper*, nr. W95/13.

Beth-Marom, R. (1982), How Probable is Probable? A Numerical Translation of Verbal Probability Expressions, *Journal of Forecasting*, 1, pp. 257-269.

Carroll, C.D. (1992), The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence, *Brooking Papers on Economic Activity*, 2, pp. 61-156.

Carroll, C.D. (1994), How does future income affect current consumption?, *Quarterly Journal of Economics*, 109, pp. 111-147.

Das, M. and A. v. Soest (1996), A Panel Data Model for Subjective Information on Household Income Growth, *CentER Discussion Paper*, nr. 9675, Tilburg University.

Das, M. and A. v. Soest (1997), Expected and Realized Income Changes: Evidence from the Dutch Socio-Economic Panel, *Journal of Economic Behavior and Organization*, forthcoming.

Dominitz, J. and C. Manski (1994), Using Expectations Data to Study Subjective Income Expectations, *NBER Working Paper*, nr. 4937.

Guiso, L., T. Jappelli, and D. Terlizzese (1992), Earnings uncertainty and precautionary saving, *Journal of Monetary Economics*, 30, pp. 307-337.

Guiso, L., T. Jappelli, and D. Terlizzese (1996), Income Risk, Borrowing Constraints and Portfolio Choice, *American Economic Review*, 86, pp. 158-172.

Härdle, W. and O. Linton (1994), Applied nonparametric methods, in R. Engle and D. McFadden (eds), *Handbook of Econometrics*, Vol. IV, North-Holland, Amsterdam, pp. 2297-2341.

Kadane, J.B. (1990), Comment: Codifying Chance, *Statistical Science*, 5, pp. 18-20.

Kimball, M.S. (1990), Precautionary Saving in the Small and in the Large, *Econometrica*, 58, pp. 53-73.

Kimball, M.S. (1993), Standard Risk Aversion, *Econometrica*, 61, pp. 589-611.

Lusardi, A. (1993), Precautionary Saving and Subjective Earnings Variance, *VSB Progress Report*, nr. 16, CentER for Economic Research, Tilburg University.

Mosteller, F. and C. Youtz (1990), Quantifying Probabilistic Expressions (with comments), *Statistical Science*, 5, pp. 2-34.

Reagan, R.T., F. Mosteller, and C. Youtz (1989), Quantitative Meanings of Verbal Probability Expressions, *Journal of Applied Psychology*, 74, pp. 433-442.

Skinner, J. (1988), Risky Income, Life Cycle Consumption and Precautionary Savings, *Journal of Monetary Economics*, 22, pp. 237-255.

Wallsten, T.S., D. Budescu, A. Rapoport, R. Zwick, and B. Forsyth (1986), Measuring the Vague Meanings of Probability Terms, *Journal of Experimental Psychology*, 115, pp. 348-365.

Zeldes, S. (1989), Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence, *Quarterly Journal of Economics*, 104, pp. 275-298.